INFORMATION AND CONTROL 41, 214–231 (1979)

Measures of Information with the Branching Property over a Graph and Their Representations[1]

C. T. NG

Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada

Let S be a set and let μ be a mapping which assigns to each finite complete probability distribution P on S a real value $\mu(P)$. We refer to such a mapping μ as a measure of information in a broad sense. To each measure μ there is associated a simple graph G on S which indicates the extent μ can be localized. It is proved that each connected component of G is either a complete graph, an arc or a polygon. Explicit representation of μ based on the structure of G is given. In particular if G is the complete graph on S, then μ can be represented as a sum.

1. INTRODUCTION

Let S be a finite set with at least four elements, and let m be a positive integer. Let $J = [0, 1]^m$ be the unit rectangle in the m-dimensional euclidean space R^m. An m-dimensional complete probability distribution P on S is a mapping of S into J such that $\sum_{s \in S} P(s) = 1$. Here 1 denotes the m-tuple in J whose coordinates are one. The set of all m-dimensional complete probability distributions on S will be denoted by $\mathscr{P}$.

A measure of information $\mu: \mathscr{P} \to R$ is a mapping which assigns to each distribution $P \in \mathscr{P}$ a value $\mu(P) \in R$.

Let a, b be distinct elements of S. The simple transfer $T_{\langle a,b\rangle}$ from a to b is the operator on $\mathscr{P}$ defined by

$$(T_{\langle a,b\rangle}P)(s) = \begin{cases} P(s) & \text{if } s \neq a, b \\ 0 & \text{if } s = a \\ P(a) + P(b) & \text{if } s = b \end{cases}$$

for all $P \in \mathscr{P}$. We call $T_{\langle a,b\rangle}P$ the accumulation of P from a to b under $T_{\langle a,b\rangle}$, and P a branching of $T_{\langle a,b\rangle}P$.

A measure (of information) $\mu: \mathscr{P} \to R$ is said to have the branching property from a to b ($a \neq b$) if the difference between $\mu(P)$ and $\mu(T_{\langle a,b\rangle}P)$ depends only on the restrictions of P and $T_{\langle a,b\rangle}P$ to $\{a, b\}$, i.e. there exists a functional Δ such that

$$\mu(P) - \mu(T_{\langle a,b\rangle}P) = \Delta(P|_{\{a,b\}}, T_{\langle a,b\rangle}P|_{\{a,b\}})$$

[1] Research supported by Canadian NRC Grant A8212.

0019-9958/79/050214-18$2.00/0

for all $P \in \mathscr{P}$. Since $(P|_{\{a,b\}}, T_{\langle a,b\rangle}P|_{\{a,b\}})$ is determined by the four-tuple $(a, b, P(a), P(b))$ we may write the above branching formula as

$$\mu(P) - \mu(T_{\langle a,b\rangle}P) = \Delta_{ab}(P(a), P(b)) \tag{B}$$

for all $P \in \mathscr{P}$. The function Δ_{ab} thus monitors the variation of μ as P is being accumulated from a to b. We refer to Δ_{ab} as a branching function.

It is observed that the branching property is not transitive. A measure (of information) may have the branching property from a to b and from b to c, and yet not from a to c. Nevertheless the notion is symmetric.

THEOREM 1.1. *If a measure of information μ has the branching property from a to b, then it has the branching property from b to a. Furthermore, the two branching functions Δ_{ab} and Δ_{ba} are related by*

$$\Delta_{ba}(y, x) = \Delta_{ab}(x, y) - \Delta_{ab}(x + y, 0) \tag{1.1.1}$$

for all $(x, y) \in D_2 := \{(x_1, x_2) \mid x_1, x_2 \in J \text{ with } x_1 + x_2 \in J\}$.

Proof. Let μ be branching from a to b with branching function Δ_{ab}, and let $P \in \mathscr{P}$ be given. Since $T_{\langle a,b\rangle} = T_{\langle a,b\rangle}T_{\langle b,a\rangle}$ as operators on $\mathscr{P}$, we have $T_{\langle a,b\rangle}P = T_{\langle a,b\rangle}T_{\langle b,a\rangle}P$ and thus

$$\mu(T_{\langle a,b\rangle}P) = \mu(T_{\langle a,b\rangle}T_{\langle b,a\rangle}P).$$

This equality and the branching formula (B) imply

$$\begin{aligned}\mu(P) - \mu(T_{\langle b,a\rangle}P) &= [\mu(P) - \mu(T_{\langle a,b\rangle}P)] - [\mu(T_{\langle b,a\rangle}P) - \mu(T_{\langle a,b\rangle}T_{\langle b,a\rangle}P)]\\ &= \Delta_{ab}(P(a), P(b)) - \Delta_{ab}(T_{\langle b,a\rangle}P(a), T_{\langle b,a\rangle}P(b))\\ &= \Delta_{ab}(P(a), P(b)) - \Delta_{ab}(P(a) + P(b), 0).\end{aligned}$$

Thus the measure μ has the branching property from b to a with the branching function Δ_{ba} given by

$$\Delta_{ba}(P(b), P(a)) := \Delta_{ab}(P(a), P(b)) - \Delta_{ab}(P(a) + P(b), 0)$$

for all $P \in \mathscr{P}$, which is the asserted relation (1.1.1).

Theorem 1.1 prompts the following:

DEFINITION 1.2. A measure (of information) μ is said to have the branching property between two distinct elements a and b of S if it has the branching property from a to b and from b to a. It is branching over a simple graph $G = (S, E)$, or G-branching if it has the branching property between the two ends of each edge in E. It follows that if $\{G_i\}$ is a family of simple graphs on S and μ is G_i-branching for all G_i, then μ is branching over the union of the G_i's. Thus to each measure μ there is associated a unique maximal simple graph G over which μ is G-branching. We refer to this maximal graph as the branching graph of μ.

In this paper we give the explicit representations of measures that are branching over a given simple graph. We also give the general structure of the branching graph of a given measure. The main results are summarized in Section 4.

In Section 2 we provide the general solution of a functional equation. The result is applied in Section 3 to give the explicit form of the branching functions which in turn give the representation of μ.

Terminologies and results concerning graphs are taken mainly from Tutte (1966). Representation for measures that are symmetric, expansible and branching is given in Ng (1974). Its application to generalized (non-probabilistic) measures of information is given in Forte and Ng (1974). Extensive literature on the subject can be found in Aczél and Daróczy (1975). A different notion of the branching property is originated and the forms of such measures are given in Forte and Bortone (1977). Relevant works on the functional equation treated in Section 2 can be found in Aczél (1965) and Hosszú (1963).

2. On the General Solution of a Functional Equation

Definition 2.1. For each positive integer n, we define

$$D_n := \left\{(x_1, x_2, \ldots, x_n) \;\middle|\; x_i \in J \text{ with } \sum_{i=1}^{n} x_i \in J\right\}.$$

The following theorem is proved in Ng (1974).

Theorem 2.2. *A function F: $D_2 \to R$ satisfies the functional equations*

$$\begin{aligned} F(x, y) &= F(y, x), \\ F(x, 0) &= 0, \\ F(x, y) + F(x + y, z) &= F(x, y + z) + F(y, z) \end{aligned} \tag{2.2.1}$$

for all $(x, y, z) \in D_3$ if, and only if, there exists a function f: $J \to R$ with initial value $f(0) = 0$ representing F through

$$F(x, y) = f(x) + f(y) - f(x + y) \tag{2.2.2}$$

for all $(x, y) \in D_2$.

The next theorem is obtained by using the above and some standard reduction technique in the methods of functional equations.

Theorem 2.3. *A system of functions F_1, F_2, F_3, F_4: $D_2 \to R$ satisfies the functional equation*

$$F_1(x, y) + F_2(x + y, z) = F_3(x, y + z) + F_4(y, z) \tag{2.3.1}$$

for all $(x, y, z) \in D_3$ *if, and only if, there exist a function* f *with initial value* $f(0) = 0$ *and a skew-symmetric bi-additive function* ψ *such that the system has the following representation for all* $(x, y) \in D_2$:

$$\begin{cases} F_1(x, y) = f(x) + f(y) - f(x+y) + \psi(x, y) + F_3(x, 0) \\ \qquad - F_3(0, 0) + F_2(y, 0) - F_2(x+y, 0) + F_1(0, y), \\ F_2(x, y) = f(x) + f(y) - f(x+y) + \psi(x, y) \\ \qquad + F_2(x, 0) + F_2(0, y) - F_2(0, 0), \\ F_3(x, y) = f(x) + f(y) - f(x+y) + \psi(x, y) \\ \qquad + F_3(x, 0) + F_3(0, y) - F_3(0, 0), \\ F_4(x, y) = f(x) + f(y) - f(x+y) + \psi(x, y) - F_3(0, x+y) \\ \qquad + F_2(x, 0) + F_2(0, y) - F_2(0, 0) + F_1(0, x). \end{cases} \tag{2.3.2}$$

3. Representation Theorems for the Branching Functions

Let $G = (S, E)$ be a simple graph and let $u, v \in S$ be adjacent vertices in G. We will denote by $[u, v]$ the unique edge in E with distinct ends u and v, and by $\langle u, v\rangle$ the path of length one traversing the edge $[u, v]$ from its origin u to its terminus v. Let $\Gamma = \langle s_0, e_1, s_1, e_2, \ldots, e_n, s_n\rangle$ be a path in G. Since G is simple, the edges occurring in Γ are uniquely determined by their ends and we may denote Γ unambiguously by $\langle s_0, s_1, \ldots, s_n\rangle$.

Let $\mu: \mathscr{P} \to R$ be a measure of information having the branching property over $G = (S, E)$, and let $\Delta_{ab}: D_2 \to R$ be the branching function of μ along the path $\langle a, b\rangle$ for each $[a, b] \in E$. We shall study some properties of the family $\{\Delta_{ab} \mid [a, b] \in E\}$ and give some representation theorems.

Let $\Gamma = \langle a_0, a_1, \ldots, a_n\rangle$ be a non-degenerate path in G. We define the transfer operator T_Γ on $\mathscr{P}$ along Γ by

$$T_\Gamma := T_{\langle a_{n-1}, a_n\rangle} T_{\langle a_{n-2}, a_{n-1}\rangle} \cdots T_{\langle a_0, a_1\rangle}, \tag{3.1}$$

which is a composition of the simple transfers along the segments of Γ in the order they occur in Γ. We call $T_\Gamma P$ the accumulation of P along Γ under the transfer T_Γ. It follows from a straightforward verification that $T_\Gamma P$ depends only on P, the terminus a_n of Γ and the set $V(\Gamma)$ of vertices occurring in Γ. This simple observation leads to the following:

Theorem 3.2. *Let* $\Gamma = \langle a_0, a_1, \ldots, a_n\rangle$ *be a non-degenerate path in* G. *Then the accumulation* $T_\Gamma P$ *of* $P \in \mathscr{P}$ *under the transfer* T_Γ *is given by*

$$T_\Gamma P(s) = \begin{cases} P(s) & \text{if } s \notin V(\Gamma), \\ 0 & \text{if } s \in V(\Gamma) \text{ and } s \neq a_n, \\ P(V(\Gamma)) & \text{if } s = a_n, \end{cases} \tag{3.2.1}$$

where $P(V(\Gamma)) := \sum_{a \in V(\Gamma)} P(a)$. *Let* $\Gamma' = \langle b_0, b_1, ..., b_m \rangle$ *be a non-degenerate path in* G, *then*

$$T_\Gamma = T_{\Gamma'} \text{ if, and only if, } a_n = b_m \text{ and } V(\Gamma) = V(\Gamma'). \tag{3.2.2}$$

If we calculate the variation of μ while P is accumulated under T_Γ, using the branching formula (B) in the order indicated by (3.1), we get

$$\begin{aligned}
&\mu(P) - \mu(T_\Gamma P) \\
&\quad = \mu(P) - \mu(T_{\langle a_0, a_1 \rangle} P) + \sum_{i=1}^{n-1} \mu(T_{\langle a_0, a_1, ..., a_i \rangle} P) - \mu(T_{\langle a_0, a_1, ..., a_{i+1} \rangle} P) \\
&\quad = \Delta_{a_0 a_1}(P(a_0), P(a_1)) + \sum_{i=1}^{n-1} \Delta_{a_i a_{i+1}}(T_{\langle a_0, a_1, ..., a_i \rangle} P(a_i), T_{\langle a_0, a_1, ..., a_i \rangle} P(a_{i+1})).
\end{aligned}$$

By introducing the definition

$$\begin{aligned}
\Delta_\Gamma(P) := {}& \Delta_{a_0 a_1}(P(a_0), P(a_1)) \\
& + \sum_{i=1}^{n-1} \Delta_{a_i a_{i+1}}(T_{\langle a_0, a_1, ..., a_i \rangle} P(a_i), T_{\langle a_0, a_1, ..., a_i \rangle} P(a_{i+1}))
\end{aligned} \tag{3.3}$$

for all $\Gamma = \langle a_0, a_1, ..., a_n \rangle$ $(n \geqslant 1)$ and $P \in \mathscr{P}$, the above equality can be rewritten as

$$\mu(P) - \mu(T_\Gamma P) = \Delta_\Gamma(P) \tag{3.4}$$

for all non-degenerate paths Γ in G and for all $P \in \mathscr{P}$. Intuitively speaking, Δ_Γ measures the variation of μ as P is accumulated along Γ.

The structure of the family $\{\Delta_{ab}\}$ of branching functions will be fully examined in what follows.

THEOREM 3.5. *The family* $\{\Delta_{ab} \mid [a, b] \in E\}$ *satisfies the boundary properties*

$$\Delta_{ab}(0, x) = 0 \qquad \text{for all} \quad x \in J. \tag{3.5.1}$$

Proof. For any $P \in \mathscr{P}$ with $P(a) = 0$ we have $T_{\langle a, b \rangle} P = P$ and thus, from the branching formula (B), $\Delta_{ab}(0, P(b)) = 0$. Since we assume Card $S \geqslant 4$, $P(b)$ takes on all possible values in J.

THEOREM 3.6. *Let* $[a, c]$, $[b, c] \in E$ *be a pair of distinct edges both incident at* $c \in S$. *Then there exist a function* $f\colon J \to R$ *with initial value* $f(0) = 0$ *and a skew-symmetric bi-additive function* $\psi\colon D_2 \to R$ *such that*

$$\begin{cases} \Delta_{ac}(x, y) = f(x) + f(y) - f(x + y) + \psi(x, y) + \Delta_{ac}(x, 0), \\ \Delta_{bc}(x, y) = f(x) + f(y) - f(x + y) - \psi(x, y) + \Delta_{bc}(x, 0) \end{cases} \tag{3.6.1}$$

for all $(x, y) \in D_2$.

Proof. Let $[a, c], [b, c] \in E$ be distinct edges incident at $c \in S$. Since the simple transfers $T_{\langle a,c\rangle}$ and $T_{\langle b,c\rangle}$ commute in the sense that $T_{\langle b,c\rangle}T_{\langle a,c\rangle} = T_{\langle a,c\rangle}T_{\langle b,c\rangle}$, it follows that $\mu(P) - \mu(T_{\langle b,c\rangle}T_{\langle a,c\rangle}P) = \mu(P) - \mu(T_{\langle a,c\rangle}T_{\langle b,c\rangle}P)$ for all $P \in \mathscr{P}$. A calculation of both sides of this equality using the branching formula (B) leads to $\mu(P) - \mu(T_{\langle b,c\rangle}T_{\langle a,c\rangle}P) = [\mu(P) - \mu(T_{\langle a,c\rangle}P)] + [\mu(T_{\langle a,c\rangle}P) - \mu(T_{\langle b,c\rangle}T_{\langle a,c\rangle}P)] = \Delta_{ac}(P(a), P(c)) + \Delta_{bc}(P(b), P(a) + P(c))$, and similarly $\mu(P) - \mu(T_{\langle a,c\rangle}T_{\langle b,c\rangle}P) = \Delta_{bc}(P(b), P(c)) + \Delta_{ac}(P(a), P(b) + P(c))$. By letting $P(a) = x$, $P(b) = z$ and $P(c) = y$ and since Card $S \geqslant 4$ we get

$$\Delta_{ac}(x, y) + \Delta_{bc}(z, x + y) = \Delta_{bc}(z, y) + \Delta_{ac}(x, z + y)$$

for all $(x, y, z) \in D_3$. We now apply Theorem 2.3 with $F_1 = F_3 = \Delta_{ac}$ and $F_2 = F_4 = \Delta_{bc}^t$ (the transpose of Δ_{bc}) and obtain representations of Δ_{ac} and Δ_{bc}^t. Taking the boundary properties (3.5.1) into our consideration we get (3.6.1). This completes the proof.

DEFINITION 3.7. Let $\mathscr{F}_0 := \{f\colon J \to R \mid f(0) = 0\}$ be the set of real-valued functions defined on J with initial value $f(0) = 0$. Let $\mathscr{A} := \{f\colon J \to R \mid f(x + y) = f(x) + f(y)$ for all $(x, y) \in D_2\} \subset \mathscr{F}_0$ be the set of additive functions on J. Two functions $f, g \in \mathscr{F}_0$ are said to be $\mathscr{A}$-congruent if $f - g \in \mathscr{A}$, and will be denoted by $f \simeq g \bmod \mathscr{A}$. The relation of being $\mathscr{A}$-congruent is an equivalence relation on $\mathscr{F}_0$. Each triple (f, g, ψ) with $f, g \in \mathscr{F}_0$ and $\psi\colon D_2 \to R$ skew-symmetric and bi-additive generates a two-place function $\Phi(f, g, \psi)$ on D_2 defined by

$$\Phi(f, g, \psi)(x, y) = f(x) + f(y) - f(x + y) + \psi(x, y) + g(x) \tag{3.7.1}$$

for all $(x, y) \in D_2$. A two-place function is said to be Φ-decomposite if it can be so generated. It is apparent that

$$\Phi(f_1, g_1, \psi_1) = \Phi(f_2, g_2, \psi_2) \quad \text{iff} \quad f_1 \simeq f_2 \bmod \mathscr{A}, \\ g_1 = g_2 \quad \text{and} \quad \psi_1 = \psi_2 . \tag{3.7.2}$$

THEOREM 3.8. *For each edge $[a, b] \in E$, the branching function Δ_{ab} is Φ-decomposite if, and only if, the branching function Δ_{ba} is Φ-decomposite. Furthermore,*

$$\Delta_{ab} = \Phi(f, g, \psi) \quad \textit{iff} \quad \Delta_{ba} = \Phi(f + g, -g, -\psi). \tag{3.8.1}$$

Proof. It follows from Theorem 1.1.

THEOREM 3.9. *Suppose $H = (V_H, E_H)$ is a connected subgraph of G having at least two edges. Then there exist functions f_b, g_{ab}, ψ_{ab} such that the branching functions Δ_{ab} are represented by*

$$\Delta_{ab} = \Phi(f_b, g_{ab}, \psi_{ab}) \tag{3.9.1}$$

for all $a, b \in V_H$ *with* $[a, b] \in E_H$. *Moreover, the functions in* $\{\psi_{ab}\}$ *satisfy*

$$\psi_{ab} = \psi_{bc} \qquad \text{and} \qquad \psi_{ab} = -\psi_{ba} \tag{3.9.2}$$

for all distinct vertices $a, b, c \in V_H$ *with* $[a, b], [b, c] \in E_H$.

Proof. Since H is connected and has at least two edges, no edge of H is isolated. Thus for each edge $[a, b] \in E_H$, either a or b is of valency larger than one, and so by Theorem 3.6 either Δ_{ba} or Δ_{ab} is Φ-decomposite. It follows from Theorem 3.8 that both Δ_{ab} and Δ_{ba} are Φ-decomposite.

Hence, to each $a, b \in V_H$ with $[a, b] \in E_H$ there correspond functions f_{ab}, g_{ab} and ψ_{ab} such that

$$\Delta_{ab} = \Phi(f_{ab}, g_{ab}, \psi_{ab}). \tag{3.9.3}$$

If we compare it with (3.6.1) using the uniqueness property (3.7.2) of Φ-decompositions we get, in particular,

$$f_{ab} \simeq f_{a'b} \bmod \mathscr{A}$$

for all $a, a', b \in V_H$ with $[a, b], [a', b] \in E_H$. Thus to each $b \in V_H$ we may choose a function $f_b \in \mathscr{F}_0$ such that

$$f_{ab} \simeq f_b \bmod \mathscr{A}$$

for all $a \in V_H$ with $[a, b] \in E_H$. With this family $\{f_b\}$ we can rewrite (3.9.3) as (3.9.1). The above comparison also leads to

$$\psi_{ab} = -\psi_{a'b} \tag{3.9.4}$$

for all distinct $a, a', b \in V_H$ with $[a, b], [a', b] \in E_H$. By comparing (3.9.3) with (3.8.1) using (3.7.2) we get, in particular,

$$\psi_{ab} = -\psi_{ba} \tag{3.9.5}$$

for all $a, b \in V_H$ with $[a, b] \in E_H$. The relations (3.9.4) and (3.9.5) among the ψ_{ab}'s are equivalent to (3.9.2). This completes the proof.

COROLLARY 3.10. *Let* H *be a connected subgraph of* G. *Suppose there exists a vertex* $a_0 \in V_H$ *with* $\mathrm{val}(H, a_0) \geqslant 3$. *Then there exist functions* f_b, $g_{ab} \in \mathscr{F}_0$ *such that the branching functions are represented by*

$$\Delta_{ab}(x, y) = f_b(x) + f_b(y) - f_b(x + y) + g_{ab}(x) \tag{3.10.1}$$

for all $(x, y) \in D_2$, *and all* $a, b \in V_H$ *with* $[a, b] \in E_H$.

Proof. Let $\{f_b, g_{ab}, \psi_{ab}\}$ be a family of functions representing the branching functions $\{\Delta_{ab}\}$ throught (3.9.1), while the family $\{\psi_{ab}\}$ satisfies (3.9.2).

Let $a_0 \in V_H$ be of valency exceeding two and let a be any given vertex adjacent to a_0 in H. There exists distinct vertices $b, c \in V_H$ different from a and adjacent to a_0. The relation (3.9.2) implies

$$\psi_{aa_0} = \psi_{a_0b}\,, \qquad \psi_{ba_0} = \psi_{a_0c}\,, \qquad \psi_{ca_0} = \psi_{a_0a}$$

$$\psi_{aa_0} = -\psi_{a_0a}\,, \qquad \psi_{ba_0} = -\psi_{a_0b}\,, \qquad \psi_{ca_0} = -\psi_{a_0c}$$

and thus it follows that all six functions vanish; in particular, $\psi_{aa_0} = \psi_{a_0a} = 0$. Since a can be any vertex in H adjacent to a_0, the ψ terms of the branching functions along all edges incident at a_0 vanish.

In view of (3.9.2) again, the vanishing of ψ_{ab} for an edge $[a, b] \in E_H$ implies that $\psi_{a'b'}$ also vanishes for any edge $[a', b']$ adjacent to $[a, b]$. This, together with the vanishing of the ψ terms of the branching functions along edges incident at a_0 and the connectedness of H, implies

$$\psi_{ab} = 0 \tag{3.10.2}$$

for all edges $[a, b] \in E_H$. The representation (3.10.1) now follows from (3.9.1) and (3.10.2). This proves the corollary.

Further analysis on the family $\{f_b\,, g_{ab}\}$ occurring in the above corollary leads to the following

THEOREM 3.11. *Let H be a connected subgraph of G. Suppose there exists a vertex in V_H with valency at least* 3. *Then there exist functions $\alpha_a \in \mathscr{F}_0$ such that the branching functions are represented by*

$$\Delta_{ab}(x, y) = \alpha_a(x) + \alpha_b(y) - \alpha_b(x + y) \tag{3.11.1}$$

for all $(x, y) \in D_2$ and all $a, b \in V_H$ with $[a, b] \in E_H$.

Proof. By Corollary 3.10 there exist functions $f_b\,, g_{ab} \in \mathscr{F}_0$ such that equation (3.10.1) holds. We shall prove the representation (3.11.1) by induction.

Consider the class $\mathscr{S}$ of all pairs (K, α), where $K = (V_K\,, E_K)$ is a subgraph of H with $E_K \neq \varnothing$ and α: $V_K \to \mathscr{F}_0$, satisfying the following hypotheses:

1° K is connected,

2° $k, k' \in V_K$ and $[k, k'] \in E_H$ imply $[k, k'] \in E_K$,

3° To each vertex $k \in V_K$, α_k is a function in $\mathscr{F}_0$. Moreover, the branching functions along edges of K are represented by

$$\Delta_{kk'}(x, y) = \alpha_k(x) + \alpha_{k'}(y) - \alpha_{k'}(x + y) \tag{3.11.2}$$

for all $(x, y) \in D_2$ and $k, k' \in V_K$ with $[k, k'] \in E_K$.

The class $\mathscr{S}$ is in fact non-empty. To see this, let $a, b \in V_H$ with $[a, b] \in E_H$

be arbitrary but fixed. Let K be the subgraph of H with vertices $V_K := \{a, b\}$ and edge set $E_K := \{[a, b]\}$. Let $\alpha\colon V_K \to \mathscr{F}_0$ be defined by

$$\alpha_b := f_b\,, \qquad \alpha_a := \alpha_b + g_{ab}\,.$$

It follows from (3.10.1) that

$$\begin{aligned} \Delta_{ab}(x, y) &= f_b(x) + f_b(y) - f_b(x + y) + g_{ab}(x) \\ &= \alpha_a(x) + \alpha_b(y) - \alpha_b(x + y) \end{aligned} \tag{3.11.3}$$

for all $(x, y) \in D_2$. A calculation of Δ_{ba} using (1.1.1) and (3.11.3) gives

$$\Delta_{ba}(x, y) = \alpha_b(x) + \alpha_a(y) - \alpha_a(x + y) \tag{3.11.4}$$

for all $(x, y) \in D_2$. Hence (K, α) satisfies hypotheses 1° to 3° and is a member of $\mathscr{S}$.

The class $\mathscr{S}$ will be considered under the partial ordering: $(K_1, \alpha_1) \leqslant (K_2, \alpha_2)$ iff K_1 is a subgraph of K_2 and α_2 is an extension of α_1. Under this partial ordering, every chain in $\mathscr{S}$ has an upper bound.

Let $(K, \alpha) \in \mathscr{S}$ and let $k, k' \in V_K$ with $[k, k'] \in E_K$ be arbitrarily given. If we compare the two alternative representations of $\Delta_{kk'}$:

$$\begin{aligned} \Delta_{kk'}(x, y) &= f_{k'}(x) + f_{k'}(y) - f_{k'}(x + y) + g_{kk'}(x + y), \\ \Delta_{kk'}(x, y) &= \alpha_k(x) + \alpha_{k'}(y) - \alpha_{k'}(x + y) \end{aligned}$$

for all $(x, y) \in D_2$, by the uniqueness of Φ-decomposition we get the relations

$$g_{kk'} = \alpha_k - \alpha_{k'} \qquad \text{and} \qquad \alpha_{k'} \simeq f_{k'} \bmod \mathscr{A}, \tag{3.11.5}$$

for all $k, k' \in V_K$ with $[k, k'] \in E_K$. These observed relations will be used later in our construction of (3.11.1).

Suppose by induction we have constructed a pair $(K, \alpha) \in \mathscr{S}$. If $V_K = V_H$ we are done. Since in this case 2° implies $E_K = E_H$ and equation (3.11.2) is then the asserted (3.11.1).

Suppose $V_K \neq V_H$. Since H is connected, there exists $a \in V_H \backslash V_K$ which is H-adjacent to some vertex $k_0 \in V_K$. Let us extend V_K to $V_{\tilde{K}} := V_K \cup \{a\}$ and E_K to $E_{\tilde{K}} := E_K \cup \{[a, k] \mid [a, k] \in E_H\,, k \in K\}$ and consider the connected subgraph $\tilde{K} = (V_{\tilde{K}}, E_{\tilde{K}})$ of H which is a proper extension of K. Let us extend α to $\tilde{\alpha}\colon V_{\tilde{K}} \to \mathscr{F}_0$ by defining

$$\tilde{\alpha}_a := \alpha_{k_0} + g_{ak_0} \qquad \text{and} \qquad \tilde{\alpha}\,|_{V_k} = \alpha. \tag{3.11.6}$$

Let $k \in V_K$ with $[a, k] \in E_H$, $k \neq k_0$ be arbitrary but fixed. We will show that $\alpha_k + g_{ak} = \alpha_{k_0} + g_{ak_0}$ and thus the definition (3.11.6) of $\tilde{\alpha}_a$ is independent of the choice of k_0 in V_K.

Since K is connected, we can join k_0 to k by a simple path $\langle k_0, k_1, \ldots, k_n = k\rangle$ in K. Let $P \in \mathscr{P}$ be a distribution with $P(a) + P(k) = 1$ and vanishing at vertices other than a and k; in particular $P(k_0) = P(k_1) = \cdots = P(k_{n-1}) = 0$. A direct calculation shows that the accumulation of P along the path $\Gamma := \langle a, k_0, k_1, \ldots, k_n = k\rangle$ or along the path $\langle a, k\rangle$ happens to be equal. This implies that

$$\Delta_\Gamma(P) = \Delta_{\langle a,k\rangle}(P).$$

If we set $x := P(a)$ and $P(k) = 1 - x$ and calculate the above variations of μ, we get

$$\begin{aligned}\Delta_\Gamma(P) &= \Delta_{ak_0}(x, 0) + \Delta_{k_0k_1}(x, 0) + \cdots + \Delta_{k_{n-2}k_{n-1}}(x, 0) + \Delta_{k_{n-1}k}(x, 1 - x)\\ &= g_{ak_0}(x) + [\alpha_{k_0}(x) - \alpha_{k_1}(x)] + \cdots + [\alpha_{k_{n-2}}(x) - \alpha_{k_{n-1}}(x)]\\ &\quad + [f_k(x) + f_k(1 - x) - f_k(1) + g_{k_{n-1}k}(x)]\\ &= g_{ak_0}(x) + \alpha_{k_0}(x) - \alpha_{k_{n-1}}(x) + [f_k(x) + f_k(1 - x) - f_k(1) + g_{k_{n-1}k}(x)]\end{aligned}$$

and

$$\begin{aligned}\Delta_{\langle a,k\rangle}(P) &= \Delta_{ak}(x, 1 - x)\\ &= f_k(x) + f_k(1 - x) - f_k(1) + g_{ak}(x)\end{aligned}$$

and thus, after eliminating common terms, we obtain

$$g_{ak_0}(x) + \alpha_{k_0}(x) - \alpha_{k_{n-1}}(x) + g_{k_{n-1}k}(x) = g_{ak}(x).$$

Since $x \in J$ is quite arbitrary, and, due to (3.11.5), $g_{k_{n-1}k} = \alpha_{k_{n-1}} - \alpha_k$, the above equality now gives

$$\alpha_{k_0} + g_{ak_0} = \alpha_k + g_{ak}\,.$$

The choice of $k \in V_K$, $k \neq k_0$, being arbitrary, we have extended definition (3.11.6) to

$$\tilde{\alpha}_a = \alpha_k + g_{ak} \tag{3.11.7}$$

for all $k \in V_K$ with $[a, k] \in E_H$.

Let $k \in K$ with $[a, k] \in E_H$ be given. From (3.10.1) we have

$$\Delta_{ak}(x, y) = f_k(x) + f_k(y) - f_k(x + y) + g_{ak}(x)$$

for all $(x, y) \in D_2$. But we also have $f_k \simeq \alpha_k \bmod \mathscr{A}$ from the relation (3.11.5), and thus

$$\Delta_{ak}(x, y) = \alpha_k(x) + \alpha_k(y) - \alpha_k(x + y) + g_{ak}(x).$$

Since, from (3.11.7), $\alpha_k(x) + g_{ak}(x) = \tilde{\alpha}_a(x)$ and $\tilde{\alpha}$ is an extension of α, this representation of Δ_{ak} now takes the form

$$\Delta_{ak}(x, y) = \tilde{\alpha}_a(x) + \tilde{\alpha}_k(y) - \tilde{\alpha}_k(x + y) \tag{3.11.8}$$

for all $(x, y) \in D_2$. A computation of Δ_{ka} via Δ_{ak} using (1.1.1) and (3.11.8) gives

$$\Delta_{ka}(x, y) = \tilde{\alpha}_k(x) + \tilde{\alpha}_a(y) - \tilde{\alpha}_a(x + y) \tag{3.11.9}$$

for all $(x, y) \in D_2$. With both (3.11.8) and (3.11.9) for an arbitrary $k \in K$ with $[a, k] \in E_H$, we have extended the hypothetical representation (3.11.2) from the system (K, α) to the higher order system $(\tilde{K}, \tilde{\alpha})$. While it is easy to see that $(\tilde{K}, \tilde{\alpha})$ satisfies hypotheses 1° and 2°, $(\tilde{K}, \tilde{\alpha})$ is in $\mathscr{S}$ and induction continues. This proves the theorem.

Remark 3.12. The representation (3.11.1) given in Theorem 3.11 is stronger than (3.10.1) given in Corollary 3.10. The relations

$$f_a \simeq \alpha_a \bmod \mathscr{A}, \qquad g_{ab} = \alpha_a - \alpha_b \tag{3.12.1}$$

for all $a, b \in V_H$ with $[a, b] \in E_H$ provide the connection between the two representations. These relations follow from (3.11.5) where we can take K to be H after Theorem 3.11 has been established.

We now come to consider the case when H is a connected subgraph of G having at least two edges and no vertices of H are of valency larger than two. In this event H is either an arc or a polygon.

THEOREM 3.13. *Let H be an arc in G with at least two edges. Let a_0 and a_n be the ends of H and let $\Gamma = \langle a_0, a_1, \ldots, a_n\rangle$ be the unique simple path in G from a_0 to a_n defining H. Then there exist functions $\alpha_{a_i} \in \mathscr{F}_0$ and a skew-symmetric bi-additive function ψ such that the branching functions $\Delta_{a_i a_j}$ are represented by*

$$\Delta_{a_i a_j}(x, y) = \alpha_{a_i}(x) + \alpha_{a_j}(y) - \alpha_{a_j}(x + y) + \operatorname{sgn}(j - i)\,\psi(x, y) \tag{3.13.1}$$

for all $(x, y) \in D_2$ and all consecutive integers $0 \leqslant i, j \leqslant n$.

Proof. The a_i's can be distinguished by their indices, and, for simplicity, each vertex a_i will be identified throughout this proof with its index i.

By Theorem 3.9, there exist functions f_i, $g_{ij} \in \mathscr{F}_0$ and skew-symmetric bi-additive functions ψ_{ij} such that

$$\Delta_{ij} = \Phi(f_j, g_{ij}, \psi_{ij})$$

for all consecutive $0 \leqslant i, j \leqslant n$. Moreover, the functions in $\{\psi_{ij}\}$ satisfy the relation

$$\psi_{ij} = \psi_{jk}, \qquad \psi_{ij} = -\psi_{ji}$$

for all consecutive pairs (i, j) and (j, k) with $i \neq k$. This implies, while letting $\psi := \psi_{01}$,

$$\psi_{ij} = \operatorname{sgn}(j - i)\psi \tag{3.13.2}$$

for all consecutive $0 \leqslant i, j \leqslant n$. Here, $\operatorname{sgn}(\omega) = 1$ if $\omega > 0$ and $\operatorname{sgn}(\omega) = -1$ if $\omega < 0$. In this case $\operatorname{sgn}(j - i) = j - i$ as i, j are consecutive integers. Anyway, sgn is kept to designate the sign. The Φ-decomposition of Δ_{ij} is now given by

$$\Delta_{ij} = \Phi(f_j, g_{ij}, \operatorname{sgn}(j-i)\psi) \tag{3.13.3}$$

for all consecutive $0 \leqslant i, j \leqslant n$.

We now define recursively the functions $\alpha_i \in \mathscr{F}_0$ by

$$\alpha_0 = f_0, \qquad \alpha_i = \alpha_{i-1} + g_{i,i-1} \qquad (1 \leqslant i \leqslant n). \tag{3.13.4}$$

From (3.13.3) and (3.13.4) we obtain

$$\begin{aligned} \Delta_{10}(x, y) &= f_0(x) + f_0(y) - f_0(x+y) - \psi(x, y) + g_{10}(x) \\ &= \alpha_1(x) + \alpha_0(y) - \alpha_0(x+y) - \psi(x, y) \end{aligned} \tag{3.13.5}$$

for all $(x, y) \in D_2$. A calculation of Δ_{01} via (3.8.1) and (3.13.5) leads to

$$\Delta_{01}(x, y) = \alpha_0(x) + \alpha_1(y) - \alpha_1(x+y) + \psi(x, y) \tag{3.13.6}$$

for all $(x, y) \in D_2$. This proves (3.13.1) when $n = 1$. If we compare the representations (3.13.3) and (3.13.6), we also get, by (3.7.2),

$$f_1 \simeq \alpha_1 \bmod \mathscr{A}. \tag{3.13.7}$$

Suppose for induction that, for some $1 \leqslant m \leqslant n$, (3.13.1) holds for all consecutive $0 \leqslant i, j \leqslant m$ and that $f_i \simeq \alpha_i$ for all $0 \leqslant i \leqslant m$; i.e.,

$$\Delta_{ij}(x, y) = \alpha_i(x) + \alpha_j(y) - \alpha_j(x+y) + \operatorname{sgn}(j-i)\, \psi(x, y) \tag{3.13.8}$$

for all $(x, y) \in D_2$ and all consecutive $0 \leqslant i, j \leqslant m$, and

$$f_i \simeq \alpha_i \bmod \mathscr{A} \tag{3.13.9}$$

for all $0 \leqslant i \leqslant m$.

If $m = n$ we are done. Let $m < n$, and consider $\Delta_{m+1,m}$ and $\Delta_{m,m+1}$. By (3.13.3) we have

$$\Delta_{m+1,m} = \Phi(f_m, g_{m+1,m}, -\psi),$$

and by (3.13.9) we have $f_m \simeq \alpha_m \bmod \mathscr{A}$. Thus, by (3.7.2) and (3.13.4), we get

$$\begin{aligned} \Delta_{m+1,m}(x, y) &= \Phi(\alpha_m, g_{m+1,m}, -\psi)(x, y) \\ &= \alpha_m(x) + \alpha_m(y) - \alpha_m(x+y) - \psi(x, y) + g_{m+1,m}(x) \\ &= \alpha_{m+1}(x) + \alpha_m(y) - \alpha_m(x+y) - \psi(x, y) \end{aligned} \tag{3.13.10}$$

for all $(x, y) \in D_2$. A calculation of $\Delta_{m,m+1}$ via (3.8.1) and $\Delta_{m+1,m}$ gives

$$\Delta_{m,m+1}(x, y) = \alpha_m(x) + \alpha_{m+1}(y) - \alpha_{m+1}(x+y) + \psi(x, y) \tag{3.13.11}$$

for all $(x, y) \in D_2$. A comparison of the Φ-decompositions of $\Delta_{m,m+1}$ in (3.13.3) and (3.13.11), while using (3.7.2), leads to

$$f_{m+1} \simeq \alpha_{m+1} \bmod \mathscr{A}. \tag{3.13.12}$$

Thus, in view of (3.13.10)–(3.13.12), the induction hypotheses (3.13.8)–(3.13.9) hold for $m + 1$. This proves the theorem.

Remark 3.14. Let H be an arc in G with at least two edges. Then the two decompositions of the branching functions along the edges of H given by (3.9.1) and (3.13.1) are related by

$$f_{a_i} \simeq \alpha_{a_i} \bmod \mathscr{A}, \qquad g_{a_i a_j} = \alpha_{a_i} - \alpha_{a_j}, \qquad \psi_{a_i a_j} = \operatorname{sgn}(j - i)\psi \tag{3.14.1}$$

for all vertices a_i, a_j adjacent in H. This follows from (3.7.2)

THEOREM 3.15. *Let H be a polygon in G with more than one edge, and let $[a_0, a_1]$ be an edge of H. Let $\Gamma = \langle a_0, a_1, ..., a_n, a_{n+1} = a_0\rangle$ be the unique circular path, whose origin is a_0 and whose first edge term is $[a_0, a_1]$, defining H. Then there exist functions $\alpha_{a_i} \in \mathscr{F}_0$ and a skew-symmetric bi-additive function ψ such that the branching functions along the edges of H are represented by*

$$\Delta_{a_i a_j}(x, y) = \alpha_{a_i}(x) + \alpha_{a_j}(y) - \alpha_{a_j}(x + y) + \operatorname{sgn}(j - i)\,\psi(x, y) \tag{3.15.1}$$

for all $(x, y) \in D_2$ and for all $0 \leqslant i, j \leqslant n$ with $|i - j| = 1$, and

$$\begin{aligned} \Delta_{a_n a_0}(x, y) &= \alpha_{a_n}(x) + \alpha_{a_0}(y) - \alpha_{a_0}(x + y) + \psi(x, y) - 2\psi(x, 1 - x), \\ \Delta_{a_0 a_n}(x, y) &= \alpha_{a_0}(x) + \alpha_{a_n}(y) - \alpha_{a_n}(x + y) - \psi(x, y) + 2\psi(x, 1 - x) \end{aligned} \tag{3.15.2}$$

for all $(x, y) \in D_2$. Moreover, ψ vanishes if the set $V(\Gamma)$ of vertices occuring in Γ is not the whole S, i.e.,

$$\{a_0, a_1, ..., a_n\} \neq S \quad \textit{implies} \quad \psi = 0. \tag{3.15.3}$$

Proof. We apply Theorem 3.13 to the residual arc of a_0 and a_n in H which contains the segment $\Gamma_1 = \langle a_0, a_1, ..., a_n\rangle$ of Γ. Hence there exist $\alpha_{a_i} \in \mathscr{F}_0$ and a skew-symmetric bi-additive function ψ such that

$$\Delta_{a_i a_j}(x, y) = \alpha_{a_i}(x) + \alpha_{a_j}(y) - \alpha_{a_j}(x + y) + \operatorname{sgn}(j - i)\,\psi(x, y) \tag{3.15.4}$$

for all $(x, y) \in D_2$, and for all $0 \leqslant i, j \leqslant n$ with $|i - j| = 1$. This proves (3.15.1).

Consider the path $\Gamma_2 = \langle a_n, a_0, a_1, ..., a_n\rangle$ and compare it with $\Gamma_1 = \langle a_0, a_1, ..., a_n\rangle$. Obviously $\Gamma_2 = \langle a_n, a_0\rangle\Gamma_1$ and $V(\Gamma_1) = V(\Gamma_2)$. It follows from Theorem 3.2 that $T_{\Gamma_1} = T_{\Gamma_2}$, and thus

$$\Delta_{\Gamma_1}(P) = \Delta_{\Gamma_2}(P) \tag{3.15.5}$$

for all $P \in \mathscr{P}$. If we let $x_i = P(a_i)$, $0 \leqslant i \leqslant n$, and calculate $\Delta_{\Gamma_1}(P)$ according to the definition of Δ_Γ given in (3.3) and using the representation (3.15.4), we get

$$\Delta_{\Gamma_1}(P) = -\alpha_{a_n}\left(\sum_{i=0}^{n} x_i\right) + \sum_{i=0}^{n} \alpha_{a_i}(x_i) + \sum_{0 \leqslant i < j \leqslant n} \psi(x_i, x_j).$$

Similar calculation of $\Delta_{\Gamma_2}(P)$ leads to

$$\Delta_{\Gamma_2}(P) = \Delta_{a_n a_0}(x_n, x_0) - \alpha_{a_n}\left(\sum_{i=0}^{n} x_i\right) + \alpha_{a_0}(x_n + x_0)$$
$$+ \sum_{i=1}^{n-1} \alpha_{a_i}(x_i) + \sum_{j=1}^{n-1} \psi(x_n, x_j) + \sum_{0 \leqslant i < j \leqslant n-1} \psi(x_i, x_j).$$

Thus (3.15.5) yields, after elimination of common terms,

$$\alpha_{a_0}(x_0) + \alpha_{a_n}(x_n) + \sum_{i=0}^{n-1} \psi(x_i, x_n)$$
$$= \Delta_{a_n a_0}(x_n, x_0) + \alpha_{a_0}(x_n + x_0) + \sum_{j=1}^{n-1} \psi(x_n, x_j),$$

which is equivalent to

$$\Delta_{a_n a_0}(x_n, x_0) = \alpha_{a_n}(x_n) + \alpha_{a_0}(x_0) - \alpha_{a_n}(x_n + x_0)$$
$$- 2\psi\left(x_n, \sum_{i=1}^{n-1} x_i\right) - \psi(x_n, x_0). \tag{3.15.6}$$

In the event $\{a_0, a_1, \ldots, a_n\} \neq S$, the above equation is to hold for all $(x_i) \in D_{n+1}$. If we let $x_i = 0$ for all $i = 1, 2, \ldots, n-1$, we find

$$\Delta_{a_n a_0}(x_n, x_0) = \alpha_{a_n}(x_n) + \alpha_{a_0}(x_0) - \alpha_{a_0}(x_n + x_0) - \psi(x_n, x_0) \tag{3.15.7}$$

for all $(x_n, x_0) \in D_2$. Subtracting it from (3.15.6) side by side, we also get

$$-2\psi\left(x_n, \sum_{i=1}^{n-1} x_i\right) = 0$$

for all $(x_i)_{i=0}^{n} \in D_{n+1}$, which is equivalent to

$$\psi = 0. \tag{3.15.8}$$

In this case, (3.15.7) gives

$$\Delta_{a_n a_0}(x, y) = \alpha_{a_n}(x) + \alpha_{a_0}(y) - \alpha_{a_0}(x + y) \tag{3.15.9}$$

for all $(x, y) \in D_2$. A calculation of $\Delta_{a_0 a_n}$ via (1.1.1) and (3.15.9) also gives

$$\Delta_{a_0 a_n}(x, y) = \alpha_{a_0}(x) + \alpha_{a_n}(y) - \alpha_{a_n}(x + y) \tag{3.15.10}$$

for all $(x, y) \in D_2$. With (3.15.8)–(3.15.10) we have established (3.15.2)–(3.15.3) in the event $\{a_0, a_1, ..., a_n\} \neq S$.

Suppose $\{a_0, a_1, ..., a_n\} = S$. Then (3.15.6) gives

$$\begin{aligned} \Delta_{a_n a_0}(x, y) &= \alpha_{a_n}(x) + \alpha_{a_n}(y) - \alpha_{a_0}(x + y) \\ &\quad - 2\psi(x, 1 - x) + \psi(x, y) \end{aligned} \tag{3.5.11}$$

for all $(x, y) \in D_2$. This gives the form of $\Delta_{a_n a_0}$ asserted in (3.15.2). A computation of $\Delta_{a_0 a_n}$ via (1.1.1) and (3.15.11) also gives the form of $\Delta_{a_0 a_n}$ in (3.15.2). This completes the proof.

4. Representation Theorem for Measures of Information with the Branching Property

Let S be a finite set with Card $S \geqslant 4$ and let $G = (S, E)$ be a simple graph on S.

The components $\mathscr{C} = \{C_\lambda\}$ of G are to be classified into four types:

$\mathscr{C}_1$: components which are vertex graphs,

$\mathscr{C}_2$: components which are link graphs,

$\mathscr{C}_3$: components where at least one vertex is of valency larger than 2, and components which are l-gons with $3 \leqslant l <$ Card S,

$\mathscr{C}_4$: components which are arcs of length at least 2, and components which are l-gons with $l =$ Card S.

In each component C_λ we choose and fix a vertex b_λ, and for convenience b_λ is chosen to be an end whenever C_λ is an arc. We refer to $\{b_\lambda\}$ as a base for $\mathscr{C}$.

For each component $C_\lambda \in \mathscr{C}_4$, we introduce a well-ordering $<_\lambda$ on the vertices in $V(C_\lambda)$ such that b_λ is minimal and consecutive pair of vertices under $<_\lambda$ are adjacent. In the event C_λ is an arc, the above well-ordering of $V(C_\lambda)$ is unique, and in the event C_λ is an l-gon, $l =$ Card S, there are two possible ways of choosing $<_\lambda$, depending on the orientation chosen.

With the above preparation and organization for the graph G, we are ready to give the following representation theorem.

Theorem 4.1. *Let S be a finite set with* Card $S \geqslant 4$, *and let $G = (S, E)$ be a simple graph on S. Let the components $\{C_\lambda\}$ of G be classified into $\mathscr{C}_i$, $i = 1, 2, 3, 4$. Let $\{b_\lambda\}$ be a chosen base for $\{C_\lambda\}$, and let $<_\lambda$ be a well-order on verices of C_λ, $C_\lambda \in \mathscr{C}_4$, as described in the preceding paragraphs.*

Let $\mathscr{P}$ be the space of all m-dimensional complete probability distributions on S and $\mu\colon \mathscr{P} \to R$ be a measure of information branching over the graph G. Then there exist: For each $C_\lambda \in \mathscr{C}_2$ a function $\Delta_\lambda\colon D_2 \to R$, for each $C_\lambda \in \mathscr{C}_3 \cup \mathscr{C}_4$ and $s \in V(C_\lambda)$ a function $\alpha_s\colon J \to R$, and for each $C_\lambda \in \mathscr{C}_4$ a skew-symmetric bi-additive function $\psi_\lambda\colon D_2 \to R$, such that μ has the following representation

$$\mu(P) = \mu(P^*) + \sum_\lambda \Delta_\lambda(P(a_\lambda), P(b_\lambda)) - \sum_\lambda \alpha_{b_\lambda}(P(C_\lambda))$$

$$+ \sum_s \alpha_s(P(s)) + \sum_\lambda \sum_{(s',s)} \psi_\lambda(P(s), P(s')) \tag{4.1.1}$$

for all $P \in \mathscr{P}$. Here $P(C_\lambda)$ denotes the P-weight $\sum_{s\in C_\lambda} P(s)$ of the component C_λ, and P^* is the accumulation of P from the components $\{C_\lambda\}$ to the base $\{b_\lambda\}$, defined by $P^*(s) = P(C_\lambda)$ if $s = b_\lambda$ and $P^*(s) = 0$ otherwise. The first summation is over all λ with C_λ, whose vertices are a_λ and b_λ, in $\mathscr{C}_2$. The second summation is over all λ with $C_\lambda \in \mathscr{C}_3 \cup \mathscr{C}_4$. The third summation is over all $s \in V(C_\lambda)$ with $C_\lambda \in \mathscr{C}_3 \cup \mathscr{C}_4$. The last double summation is over all λ with $C_\lambda \in \mathscr{C}_4$ and (s', s) with $s' <_\lambda s$. Moreover, the functions Δ_λ, α_s satisfy the following boundary conditions

$$\Delta_\lambda(0, \cdot) = 0, \qquad \alpha_s(0) = 0. \tag{4.1.2}$$

The converse is also true: Let Δ_λ, α_s be arbitrary functions satisfying the boundary conditions (4.1.2), *ψ_λ be arbitrary skew-symmetric bi-additive functions and μ be arbitrarily initiated for distributions P^* on S with support $\{b_\lambda\}$. Then equation* (4.1.1) *defines a measure μ which is G-branching.*

Proof. Let μ be G-branching, and let $\Delta_{ab}\colon D_2 \to R$, $[a, b] \in E$, be the branching functions associated with μ.

For each $C_\lambda \in \mathscr{C}_2$, let its two vertices be a_λ and b_λ, and let Δ_λ be $\Delta_{a_\lambda b_\lambda}$. The boundary condition (4.1.2) for Δ_λ follows from Theorem 3.5.

For each $C_\lambda \in \mathscr{C}_3$, either Theorem 3.11 or Theorem 3.15 is applicable. Hence to each $s \in V(C_\lambda)$ there corresponds a function $\alpha_s \in \mathscr{F}_0$ such that the branching functions along the edges of C_λ have the representation (3.11.1), which coincides with the representation (3.15.1) together with (3.15.3).

For each $C_\lambda \in \mathscr{C}_4$, either Theorem 3.13 or Theorem 3.15 is applicable with a simple path Γ_λ covering the vertices of C_λ in descending order of $<_\lambda$ and terminating at b_λ. There exist functions $\alpha_s \in \mathscr{F}_0$, $s \in V(C_\lambda)$, and a skew-symmetric bi-additive function ψ_λ such that the branching functions along those edges of C_λ occuring in Γ_λ have the representation (3.13.1), which coincides with (3.15.1), where the ψ-term takes a positive sign if $a_i >_\lambda a_j$ (i.e., $j > i$) and negative otherwise.

To each component $C_\lambda \in \mathscr{C}_i$, $i = 2, 3, 4$, we now choose a path (not necessarily simple) covering all vertices in $V(C_\lambda)$ and terminating at the base point b_λ. We can now calculate the variation $\mathscr{V}_\lambda$ of μ as P is accumulated along this path, using formula (3.3) and the explicit forms of the branching functions. It turns out that $\mathscr{V}_\lambda$ is given by

$$\begin{aligned} &\Delta_\lambda(P(a_\lambda), P(b_\lambda)) &&\text{if}\quad C_\lambda \in \mathscr{C}_2,\\ -&\alpha_{b_\lambda}(P(C_\lambda)) + \sum_s \alpha_s(P(s)) &&\text{if}\quad C_\lambda \in \mathscr{C}_3,\\ -&\alpha_{b_\lambda}(P(C_\lambda)) + \sum_s \alpha_s(P(s)) + \sum_{(s',s)} \psi_\lambda(P(s), P(s')) &&\text{if}\quad C_\lambda \in \mathscr{C}_4, \end{aligned}$$

where the summations are over all $s \in V(C_\lambda)$, and (s', s) with $s' <_\lambda s$ respectively. This reduction can be carried out successively from component to component of G until P is reduced to P^*, and the relation $\mu(P) = \mu(P^*) + \sum_\lambda \mathscr{V}_\lambda$ is found to be (4.1.1).

The converse is straightforward. The boundary properties (4.1.2) assure that equation (4.1.1) extends μ from distributions P^*, with support contained in the base $\{b_\lambda\}$, to $\mathscr{P}$ consistently. The measure μ so defined is indeed G-branching. This completes the proof.

COROLLARY 4.2. *A measure μ is branching over the complete simple graph on S if, and only if, it has the form*

$$\mu(P) = c + \sum \alpha_s(P(s)) \qquad (P \in \mathscr{P}),$$

where the summation is over all $s \in S$, α_s: $J \to R$ are functions satisfying the boundary conditions $\alpha_s(0) = 0$, and c is a constant.

Proof. In this case we can choose $b_0 \in S$ arbitrarily as a base point and the measure μ has the form

$$\mu(P) = \mu(P^*) - \alpha_{b_0}(1) + \sum_s \alpha_s(P(s)).$$

The term $\mu(P^*)$ is a constant, since P^* is the distribution whose weight at b_0 is 1 and does not depend on P.

COROLLARY 4.3. *Let μ be a measure (of information), and let G be its (maximal) branching graph as defined in Definition 1.2. Then all components of G that are of the type $\mathscr{C}_3$ are complete; and no component of G is an arc whose set of vertices is the whole S.*

RECEIVED: December 16, 1977; REVISED: November, 1978

REFERENCES

J. ACZÉL (1965), The general solution of two functional equations by reduction to functions additive in two variables and with the aid of Hamel bases, *Glasnik Mat.-Fiz. Astr.* **20**, 65–73.

J. ACZÉL AND Z. DARÓCZY (1975), "On Measures of Information and Their Characterizations," pp. 103–113, Academic Press, New York.

B. FORTE AND C. T. NG (1974), Entropies with the branching property, *Ann. Mat. Pura Appl.* 4, 355–373.

B. FORTE AND C. A. BORTONE (1977), Non-symmetric entropies with the branching property, *Utilitas Math.* 12, 3–23.

M. HOSSZÚ (1963), On a functional equation treated by S. Kurepa, *Glasnik Mat.-Fiz. Astr.* 18, 59–60.

C. T. NG (1974), Representation for measures of information with the branching property, *Inform. Contr.* 25, 45–56.

W. T. TUTTE (1966), Connectivity in graphs, Mathematical Expositions No. 15, Univ. of Toronto, Canada.